

OIKE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/915,694

DATE: 01/16/2002

TIME: 18:25:23

Input Set : N:\Crf3\RULE60\09915694.raw

Output Set: N:\CRF3\01162002\I915694.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Bandman, Olga

6 Corely, Neil C.

7 Shah, Purvi

9 (ii) TITLE OF INVENTION: HUMAN MITOCHONDRIAL MALATE DEHYDROGENASE

11 (iii) NUMBER OF SEQUENCES: 4

13 (iv) CORRESPONDENCE ADDRESS:

14 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.

15 (B) STREET: 3174 Porter Drive

16 (C) CITY: Palo Alto

17 (D) STATE: CA

18 (E) COUNTRY: USA

19 (F) ZIP: 94304

21 (v) COMPUTER READABLE FORM:

22 (A) MEDIUM TYPE: Diskette

23 (B) COMPUTER: IBM Compatible

24 (C) OPERATING SYSTEM: DOS

25 (D) SOFTWARE: FastSEQ for Windows Version 2.0

27 (vi) CURRENT APPLICATION DATA:

C--> 28 (A) APPLICATION NUMBER: US/09/915,694

C--> 29 (B) FILING DATE: 25-Jul-2001

30 (C) CLASSIFICATION:

32 (vii) PRIOR APPLICATION DATA:

33 (A) APPLICATION NUMBER: 08/922,957

34 (B) FILING DATE:

36 (viii) ATTORNEY/AGENT INFORMATION:

37 (A) NAME: Billings, Lucy J.

38 (B) REGISTRATION NUMBER: 36,749

39 (C) REFERENCE/DOCKET NUMBER: PF-0379 US

41 (ix) TELECOMMUNICATION INFORMATION:

42 (A) TELEPHONE: 650-855-0555

43 (B) TELEFAX: 650-845-4166

44 (C) TELEX:

47 (2) INFORMATION FOR SEQ ID NO: 1:

49 (i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 338 amino acids

51 (B) TYPE: amino acid

52 (C) STRANDEDNESS: single

53 (D) TOPOLOGY: linear

55 (vii) IMMEDIATE SOURCE:

56 (A) LIBRARY: THP1PLB01

57 (B) CLONE: 11587

59 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

61 Met Leu Ser Ala Leu Ala Arg Pro Ala Ser Ala Ala Leu Arg Arg Ser

62 1 5 10 15

63 Phe Ser Thr Ser Ala Gln Asn Asn Ala Lys Val Ala Val Leu Gly Ala

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/915,694

DATE: 01/16/2002

TIME: 18:25:23

Input Set : N:\Crif3\RULE60\09915694.raw

Output Set: N:\CRF3\01162002\I915694.raw

```

64          20          25          30
65 Ser Gly Gly Ile Gly Gln Pro Leu Ser Leu Leu Leu Lys Asn Ser Pro
66          35          40          45
67 Leu Val Ser Arg Leu Thr Leu Tyr Asp Ile Ala His Thr Pro Gly Val
68          50          55          60
69 Ala Ala Asp Leu Ser His Ile Glu Thr Lys Ala Ala Val Lys Gly Tyr
70          65          70          75          80
71 Leu Gly Pro Glu Gln Leu Pro Asp Cys Leu Lys Gly Cys Asp Val Val
72          85          90          95
73 Val Ile Pro Ala Gly Val Pro Arg Lys Pro Gly Met Thr Arg Asp Asp
74          100          105          110
75 Leu Phe Asn Thr Asn Ala Thr Ile Val Ala Thr Leu Thr Ala Ala Cys
76          115          120          125
77 Ala Gln His Cys Pro Glu Ala Met Ile Cys Val Ile Ala Asn Pro Val
78          130          135          140
79 Asn Ser Thr Ile Pro Ile Thr Ala Glu Val Phe Lys Lys His Gly Val
80          145          150          155          160
81 Tyr Asn Pro Asn Lys Ile Phe Gly Val Thr Thr Leu Asp Ile Val Arg
82          165          170          175
83 Ala Asn Thr Phe Val Ala Glu Leu Lys Gly Leu Asp Pro Ala Arg Val
84          180          185          190
85 Asn Val Pro Val Ile Gly Gly His Ala Gly Lys Thr Ile Ile Pro Leu
86          195          200          205
87 Ile Ser Gln Cys Thr Pro Lys Val Asp Phe Pro Gln Asp Gln Leu Thr
88          210          215          220
89 Ala Leu Thr Gly Arg Ile Gln Glu Ala Gly Thr Glu Val Val Lys Ala
90          225          230          235          240
91 Lys Ala Gly Ala Gly Ser Ala Thr Leu Ser Met Ala Tyr Ala Gly Ala
92          245          250          255
93 Arg Phe Val Phe Ser Leu Val Asp Ala Met Asn Gly Lys Glu Gly Val
94          260          265          270
95 Val Glu Cys Ser Phe Val Lys Ser Gln Glu Thr Glu Cys Thr Tyr Phe
96          275          280          285
97 Ser Thr Pro Leu Leu Leu Gly Lys Lys Gly Ile Glu Lys Asn Leu Gly
98          290          295          300
99 Ile Gly Lys Val Ser Ser Phe Glu Glu Lys Met Ile Ser Asp Ala Ile
100          305          310          315          320
101 Pro Glu Leu Lys Ala Ser Ile Lys Lys Gly Glu Asp Phe Val Lys Thr
102          325          330          335
103 Leu Lys

```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1334 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: THP1PLB01

(B) CLONE: 11587

RAW SEQUENCE LISTING

DATE: 01/16/2002

PATENT APPLICATION: US/09/915,694

TIME: 18:25:23

Input Set : N:\Crif3\RULE60\09915694.raw

Output Set: N:\CRF3\01162002\I915694.raw

```

118      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
120      GGCCCCAGAG AGCAGGCGCT GGGCAGTGTG GAGGTCGTTG GAGTCACTTC CGCGTCACCA      60
121      GCTCCTGTGC CTGCCAGTCG GTGCCCCCTCC CGCTCCAGCC ATGCTCTCCG CCCTCGCCCCG      120
122      GCCTGCCAGC GCTGCTCTCC GCCGCAGCTT CAGCACCTCG GCCCAGAACA ATGCTAAAGT      180
123      AGCTTGCTA GGGGCCTCTG GAGGCATCGG GCAGCCACTT TCACTTCTCC TGAAGAACAG      240
124      CCCCTTGGTG AGCCGCCTGA CCCTCTATGA TATCGCGCAC ACACCCGGAG TGGCCGCAGA      300
125      TCTGAGCCAC ATCGAGACCA AAGCCGCTGT GAAAGGCTAC CTCGGACCTG AACAGCTGCC      360
126      TGAAGTGCCTG AAAGGTTGTG ATGTGGTAGT TATTCCGGCT GGAGTCCCCA GAAAGCCAGG      420
127      CATGACCCGG GACGACCTGT TCAACACCAA TGCCACGATT GTGGCCACCC TGACCGCTGC      480
128      CTGTGCCAG CACTGCCCGG AAGCCATGAT CTGCGTCATT GCCAATCCGG TTAATTCCAC      540
129      CATCCCCATC ACAGCAGAAG TTTTCAAGAA GCATGGAGTG TACAACCCCA ACAAATCTT      600
130      CGGCGTGACG ACCCTGGACA TCGTCAGAGC CAACACCTTT GTTGCCAGAGC TGAAGGGTTT      660
131      GGATCCAGCT CGAGTCAACG TCCCTGTCAT TGGTGGCCAT GCTGGGAAGA CCATCATCCC      720
132      CCTGATCTCT CAGTGCACCC CCAAGGTGGA CTTTCCCCAG GACCAGCTGA CAGCACTCAC      780
133      TGGGCGGATC CAGGAGGCCG GCACGGAGGT GGTCAAGGCT AAAGCCGGAG CAGGCTCTGC      840
134      CACCCTCTCC ATGGCGTATG CCGGCGCCCCG CTTTGTCTTC TCCCTTGTGG ATGCAATGAA      900
135      TGGAAAGGAA GGTGTTGTGG AATGTTCCCTT CGTTAAGTCA CAGGAAACGG AATGTACCTA      960
136      CTTCTCCACA CCGCTGCTGC TTGGGAAAAA GGGCATCGAG AAGAACCCTG GCATCGGCAA      1020
137      AGTCTCCTCT TTTGAGGAGA AGATGATCTC GGATGCCATC CCCGAGCTGA AGGCCTCCAT      1080
138      CAAGAAGGGG GAAGATTTTC TGAAGACCTT GAAGTGAGCC GCTGTGACGG GTGGCCAGTT      1140
139      TCCTTAATTT ATGAAGGCAT CATGTCACG CAAAGCCGTT GCAGATAAAC TTTGTATTTT      1200
140      AATTTGCTTT GGTGATGATT ACTGTATTGA CATCATCATG CCTTCCAAAT TGTGGGTGGC      1260
141      TCTGTGGGCG CATCAATAAA AGCCGTCCTT GATTTTATTT TTCAAGGTCC CTTCTGTAAA      1320
142      TGCAAAAAAA AAAA      1334
144      (2) INFORMATION FOR SEQ ID NO: 3:
146      (i) SEQUENCE CHARACTERISTICS:
147          (A) LENGTH: 338 amino acids
148          (B) TYPE: amino acid
149          (C) STRANDEDNESS: single
150          (D) TOPOLOGY: linear
152      (vii) IMMEDIATE SOURCE:
153          (A) LIBRARY: GenBank
154          (B) CLONE: 56643
156      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
158      Met Leu Ser Ala Leu Ala Arg Pro Val Gly Ala Ala Leu Arg Arg Ser
159      1      5      10      15
160      Phe Ser Thr Ser Ala Gln Asn Asn Ala Lys Val Ala Val Leu Gly Ala
161      20      25      30
162      Ser Gly Gly Ile Gly Gln Pro Leu Ser Leu Leu Leu Lys Asn Ser Pro
163      35      40      45
164      Leu Val Ser Arg Leu Thr Leu Tyr Asp Ile Ala His Thr Pro Gly Val
165      50      55      60
166      Ala Ala Asp Leu Ser His Ile Glu Thr Arg Ala Asn Val Lys Gly Tyr
167      65      70      75      80
168      Leu Gly Pro Glu Gln Leu Pro Asp Cys Leu Lys Gly Cys Asp Val Val
169      85      90      95
170      Val Ile Pro Ala Gly Val Pro Arg Lys Pro Gly Met Thr Arg Asp Asp
171      100      105      110
172      Leu Phe Asn Thr Asn Ala Thr Ile Val Ala Thr Leu Thr Ala Ala Cys

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/915,694

DATE: 01/16/2002

TIME: 18:25:23

Input Set : N:\Crif3\RULE60\09915694.raw

Output Set: N:\CRF3\01162002\I915694.raw

```

173          115          120          125
174 Ala Gln His Cys Pro Glu Ala Met Ile Cys Ile Ile Ser Asn Pro Val
175          130          135          140
176 Asn Ser Thr Ile Pro Ile Thr Ala Glu Val Phe Lys Lys His Gly Val
177 145          150          155          160
178 Tyr Asn Pro Asn Lys Ile Phe Gly Val Thr Thr Leu Asp Ile Val Arg
179          165          170          175
180 Ala Asn Thr Phe Val Ala Glu Leu Lys Gly Leu Asp Pro Ala Arg Val
181          180          185          190
182 Asn Val Pro Val Ile Gly Gly His Ala Gly Lys Thr Ile Ile Pro Leu
183          195          200          205
184 Ile Ser Gln Cys Thr Pro Lys Val Asp Phe Pro Gln Asp Gln Leu Ala
185          210          215          220
186 Thr Leu Thr Gly Lys Ile Gln Glu Ala Gly Thr Glu Val Val Lys Ala
187 225          230          235          240
188 Lys Ala Gly Ala Gly Ser Ala Thr Leu Ser Met Ala Tyr Ala Gly Ala
189          245          250          255
190 Arg Phe Val Phe Ser Leu Val Asp Ala Met Asn Gly Lys Glu Gly Val
191          260          265          270
192 Ile Glu Cys Ser Phe Val Gln Ser Lys Glu Thr Glu Cys Thr Tyr Phe
193          275          280          285
194 Ser Thr Pro Leu Leu Leu Gly Lys Lys Gly Leu Glu Lys Asn Leu Gly
195          290          295          300
196 Ile Gly Lys Ile Thr Pro Phe Glu Glu Lys Met Ile Ala Glu Ala Ile
197 305          310          315          320
198 Pro Glu Leu Lys Ala Ser Ile Lys Lys Gly Glu Asp Phe Val Lys Asn
199          325          330          335
200 Met Lys

```

203 (2) INFORMATION FOR SEQ ID NO: 4:

205 (i) SEQUENCE CHARACTERISTICS:

206 (A) LENGTH: 298 amino acids

207 (B) TYPE: amino acid

208 (C) STRANDEDNESS: single

209 (D) TOPOLOGY: linear

211 (vii) IMMEDIATE SOURCE:

212 (A) LIBRARY: GenBank

213 (B) CLONE: 164541

215 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

217 Ser Leu Leu Leu Lys Asn Ser Pro Leu Val Ser Arg Leu Thr Leu Tyr
218 1          5          10          15
219 Asp Ile Ala His Thr Pro Gly Val Ala Ala Asp Leu Ser His Ile Glu
220          20          25          30
221 Thr Arg Ala Thr Val Lys Gly Tyr Leu Gly Pro Glu Gln Leu Pro Asp
222          35          40          45
223 Cys Leu Lys Gly Cys Asp Val Val Val Ile Pro Ala Gly Val Pro Arg
224          50          55          60
225 Lys Pro Gly Met Thr Arg Asp Asp Leu Phe Asn Thr Asn Ala Thr Met
226 65          70          75          80
227 Val Ala Thr Leu Thr Val Ala Cys Ala Gln His Cys Pro Asp Ala Met

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/915,694

DATE: 01/16/2002

TIME: 18:25:23

Input Set : N:\Crf3\RULE60\09915694.raw

Output Set: N:\CRF3\01162002\I915694.raw

228				85				90				95		
229	Ile	Cys	Ile	Ile	Ser	Asn	Pro	Val	Asn	Ser	Thr	Ile	Pro	Met
230				100					105					110
231	Glu	Val	Phe	Lys	Lys	His	Gly	Val	Tyr	Asn	Pro	Asn	Lys	Ile
232				115					120					125
233	Val	Thr	Thr	Leu	Asp	Ile	Val	Arg	Ala	Asn	Ala	Phe	Val	Ala
234				130					135					140
235	Lys	Gly	Leu	Asp	Pro	Ala	Arg	Val	Ser	Val	Pro	Val	Ile	Gly
236				145										150
237	Ala	Gly	Lys	Thr	Ile	Ile	Pro	Leu	Ile	Ser	Gln	Cys	Thr	Pro
238				165										170
239	Asp	Phe	Pro	Gln	Asp	Gln	Leu	Ser	Thr	Leu	Thr	Gly	Arg	Ile
240				180										185
241	Ala	Gly	Thr	Glu	Val	Val	Lys	Ala	Lys	Ala	Gly	Ala	Gly	Ser
242				195										200
243	Leu	Ser	Met	Ala	Tyr	Ala	Gly	Ala	Arg	Phe	Val	Phe	Ser	Leu
244				210										215
245	Ala	Met	Asn	Gly	Lys	Glu	Gly	Val	Val	Glu	Cys	Ser	Phe	Val
246				225										230
247	Gln	Glu	Thr	Asp	Cys	Pro	Tyr	Phe	Ser	Thr	Pro	Leu	Leu	Leu
248					245									250
249	Lys	Gly	Ile	Glu	Lys	Asn	Leu	Arg	Ile	Gly	Lys	Ile	Ser	Pro
250				260										265
251	Glu	Lys	Met	Ile	Ala	Glu	Ala	Ile	Pro	Glu	Leu	Lys	Ala	Ser
252				275										280
253	Lys	Gly	Glu	Glu	Phe	Val	Lys	Asn	Thr	Lys				
254				290										295

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/915,694

DATE: 01/16/2002

TIME: 18:25:24

Input Set : N:\Crf3\RULE60\09915694.raw

Output Set: N:\CRF3\01162002\I915694.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]